Selected examples for Chapter 3.

EXAMPLE 3.1

At a certain time n_o a Bernoulli process is observed to have the value $x[n_o] = +1$. It is desired to compute the distribution for the waiting time to the next '+1' in the sequence.

Suppose the time at which the next +1 occurs is at $n = n_o + l$. Then there must be $l - 1$ consecutive -1's occuring before the +1. Thus

$$\Pr[\text{time to the next } +1 \text{ is } l] = (1 - \mathrm{P})^{l-1}\mathrm{P}$$

for $l = 1, 2, 3, \ldots$. Note that the same answer would hold if $x[n_o]$ were equal to -1. This shows that *the Bernoulli process has no memory.* Further, the result is independent of the value of n_o. This again demonstrates the stationarity. $\square$

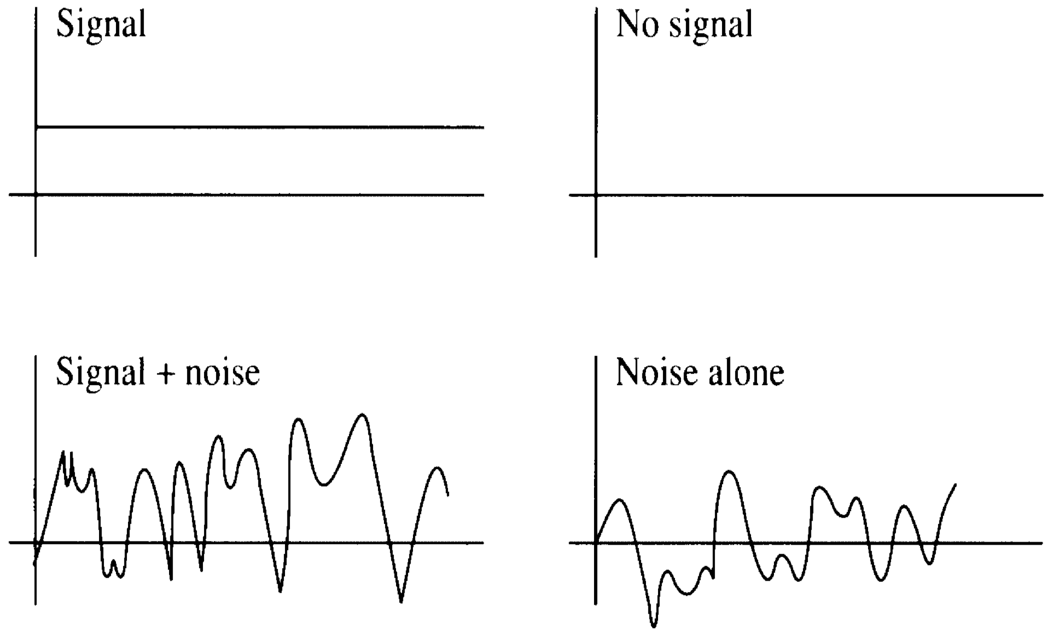
EXAMPLE 3.2

A *counting process* is defined as a random process that counts the number of positive values in a Bernoulli process after a given time n_o up to any time $n > n_o$. Taking $n_o = 0$ yields

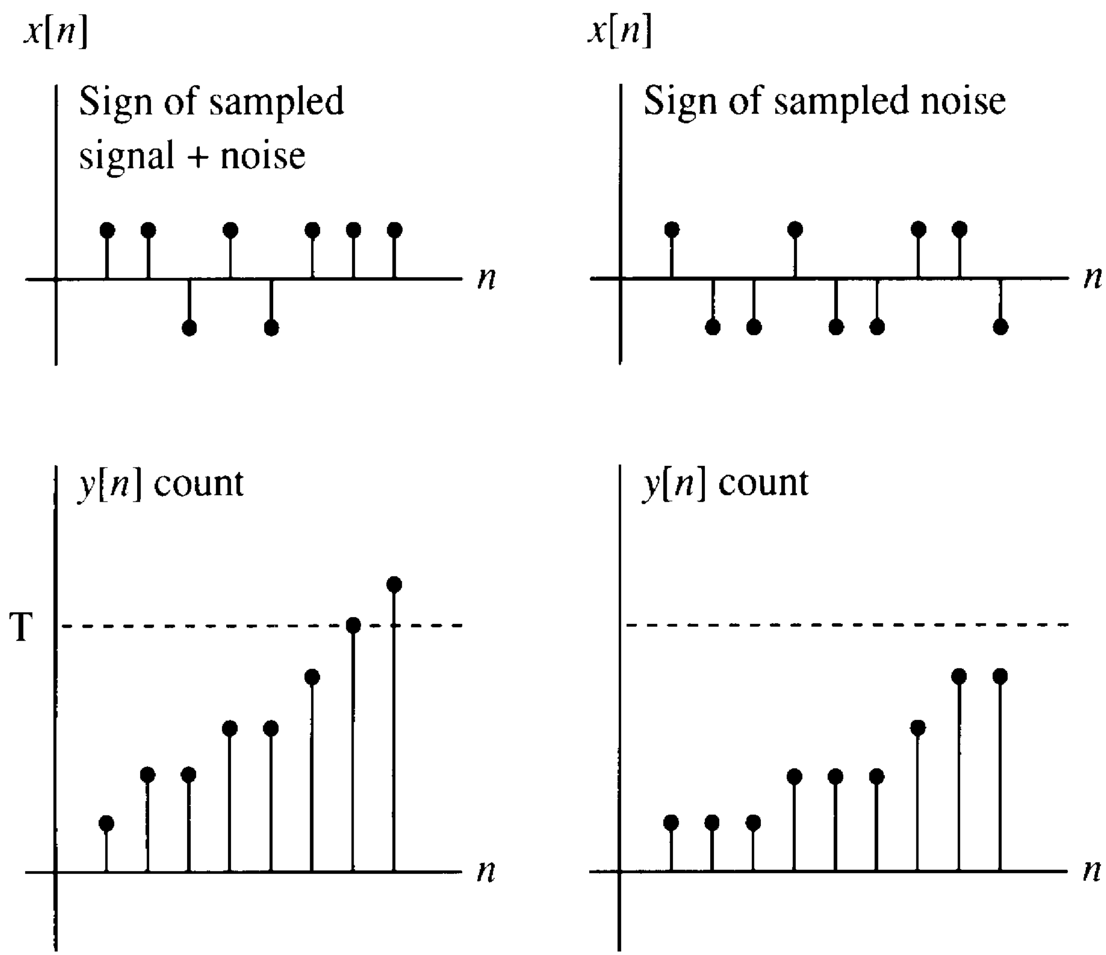
$$y[n] = \sum_{l=1}^{n} (x[l] + 1)/2$$

Each time $x[l]$ takes on a value of $+1$ the corresponding term in the sum is equal to one and each time $x[l]$ is -1, the corresponding term is zero. Thus $y[n]$ represents the number of positive values in the sequence from $l = 1$ to $l = n$.

The counting process can be used to model the evolution of observations in a nonparametric signal detection procedure known as a sign test. In this test a positive-valued signal is observed in zero-mean noise.



To give the detection procedure robust performance for noise processes with a wide variety of statistical characteristics, only the *sign* of the observations is used in the detector. Intuitively, for any fixed number of observations of the received sequence, we would expect to find that there are more positive values if the signal is present, and about equal numbers of positive and negative values if the signal is not present. The number of positive values, which can be modeled as a counting process, is compared to a threshold to make a detection decision.



Notice that the counting process looks like a set of steps occuring at random times. Observe that at time n, $y[n]$ can take on any integer values between 0 and n. To compute the probability distribution of this random process suppose that $y[n]$ is equal to some value r. This means that there are r positive values and $n-r$ negative values between $l=0$ and $l=n$. These can occur in *any* order. The number of combinations of positive and negative values is given by the binomial coefficient

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

Thus it follows that

$$\Pr[y[n]=r] = \begin{cases} \binom{n}{r}\,\mathrm{P}^r(1-\mathrm{P})^{n-r} & r=0,1,\ldots,n \\ 0 & \text{otherwise} \end{cases}$$

This is the binomial distribution. $\square$

EXAMPLE 3.7

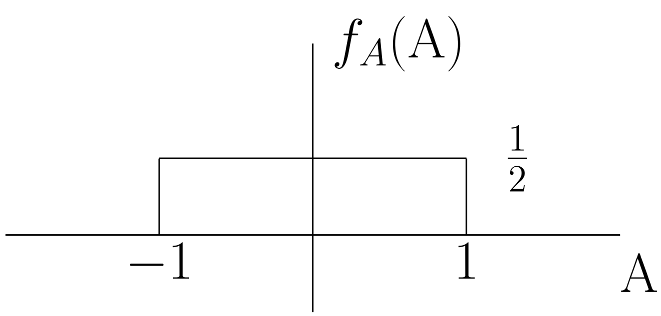
A sinusoidal random process has the form

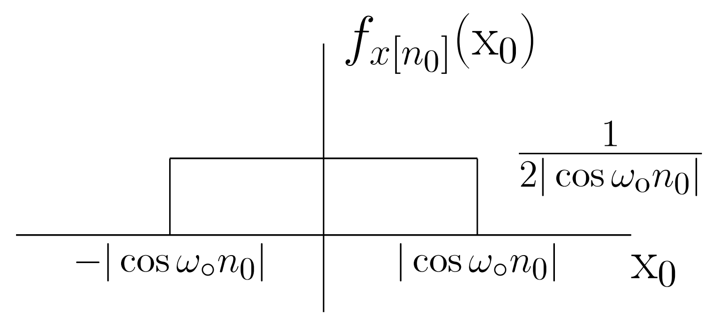
$$x[n] = A\cos\omega_o n$$

where A is a zero-mean random variable with density $f_A(\mathrm{A})$. The probability density functions for the random process can be computed as follows. For any choice of the parameter $n = n_0$ the single sample $x[n_0]$ is just a known constant $(\cos\omega_o n_0)$ times the random variable A. Therefore the first order density is simply

$$f_{x[n_0]}(\mathrm{x}_0) = \frac{1}{|\cos\omega_o n_0|} f_A\left(\frac{\mathrm{x}_0}{\cos\omega_o n_0}\right)$$

as long as $\cos\omega_o n_0 \neq 0$. For example, if A is uniformly distributed between -1 and $+1$, the densities appear as shown below:



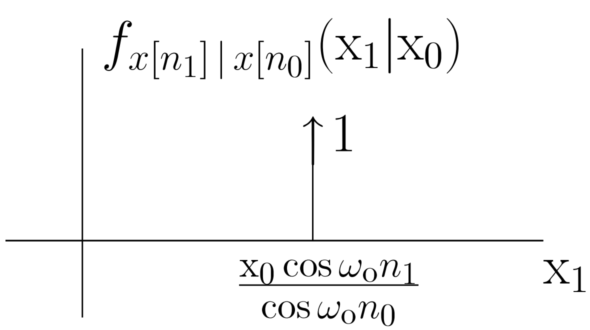


Note that as $\cos\omega_o n_0 \to 0$ the density for the sample approaches a unit impulse. This is as it should be since at a zero crossing the value of $x[n_0]$ is 0 with probability one.

To find the joint density for two samples, we first find the conditional density $f_{x[n_1]\,|\,x[n_0]}$ and multiply this by $f_{x[n_0]}$ to get $f_{x[n_0]x[n_1]}$. To find the conditional density, assume that the value $x[n_0] = \mathrm{x}_0$ was observed. Then since $A = \mathrm{x}_0/\cos\omega_o n_0$, the value of $x[n_1]$ is precisely equal to

$$\mathrm{x}_1 = \left(\frac{\mathrm{x}_0}{\cos\omega_o n_0}\right)\cos\omega_o n_1$$

In other words, given $x[n_0]$, the random variable $x[n_1]$ is known with certainty. Its conditional density is therefore an impulse as shown below.



$$f_{x[n_1]\,|\,x[n_0]}(\mathrm{x}_1|\mathrm{x}_0) = \delta_c\left(\mathrm{x}_1 - \frac{\mathrm{x}_0\cos\omega_o n_1}{\cos\omega_o n_0}\right)$$

Therefore

$$\begin{aligned} f_{x[n_0]x[n_1]}(\mathrm{x}_0, \mathrm{x}_1) &= f_{x[n_0]}(\mathrm{x}_0) f_{x[n_1]\,|\,x[n_0]}(\mathrm{x}_1|\mathrm{x}_0) \\ &= \frac{1}{|\cos\omega_\mathrm{o} n_0|} f_A\left(\frac{\mathrm{x}_0}{\cos\omega_\mathrm{o} n_0}\right) \delta_c\left(\mathrm{x}_1 - \frac{\mathrm{x}_0 \cos\omega_\mathrm{o} n_1}{\cos\omega_\mathrm{o} n_0}\right) \end{aligned}$$

Higher order densities can be computed by multiplying by additional terms. For example, to compute $f_{x[n_0]x[n_1]x[n_2]}$ multiply the above result by

$$f_{x[n_2]\,|\,x[n_0]x[n_1]}(\mathrm{x}_2|\mathrm{x}_0\mathrm{x}_1) = \delta_c\left(\mathrm{x}_2 - \frac{\mathrm{x}_1 \cos\omega_\mathrm{o} n_2}{\cos\omega_\mathrm{o} n_1}\right)$$

By continuing in this way the joint density for any number of samples can be derived.

From the foregoing result and the definition

$$f_{x[n_0],x[n_1],\ldots x[n_L]} = f_{x[n_0+k_0P],x[n_1+k_1P],\ldots x[n_L+k_LP]}$$

it can be seen that the *random process* $x[n]$ is periodic if and only if ω_o is a rational multiple of 2π, that is if ω_o is of the form $\omega_\mathrm{o} = 2\pi K/P$ where K and P are both integers. In this case the random process is periodic with period equal to $P = 2\pi K/\omega_\mathrm{o}$. $\square$

EXAMPLE 3.8

A Gaussian 'white noise' process consists of a sequence of zero-mean independent Gaussian random variables $w[n]$ with variance σ_{o}^2. Because a Gaussian random vector is defined by only its first two moments, a complete statistical description of this random process requires only being able to write the mean vector and covariance matrix for any set of samples of this random process. Since the mean is zero and the samples are independent, the white noise process has the simple characterization

$$\mathbf{m}_{\boldsymbol{w}} = \mathbf{0}$$

and

$$\mathbf{C}_{\boldsymbol{w}} = \begin{bmatrix} \sigma_{\mathrm{o}}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{\mathrm{o}}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{\mathrm{o}}^2 \end{bmatrix} = \sigma_{\mathrm{o}}^2 \mathbf{I}$$

where the dimensions of the vector and matrix are determined by the number of samples taken.

The Gaussian white noise sequence is applied to a linear filter which produces an output random process according to the difference equation

$$x[n] = \rho x[n-1] + w[n]$$

with real coefficient ρ. From previous considerations, this output random process is known to be both a Gaussian random process and a Markov process. Such a random process is sometimes called a *Gauss-Markov process*. Now assume for convenience that the process is generated from the difference equation beginning at $n_o = 0$ with initial conditions $x[-1] = 0$. Then by carrying out successive steps of the recursion the output can be written in the convolution form

$$x[n] = \sum_{k=0}^{n} w[k]\rho^{n-k}$$

or in the matrix form

$$\begin{bmatrix} x[0] \\ x[1] \\ x[2] \\ \vdots \\ x[n] \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \rho & 1 & 0 & \cdots & 0 \\ \rho^2 & \rho & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^n & \rho^{n-1} & & \cdots & 1 \end{bmatrix} \begin{bmatrix} w[0] \\ w[1] \\ w[2] \\ \vdots \\ w[n] \end{bmatrix}$$

Since this is in the form $\boldsymbol{x} = \mathbf{A}\boldsymbol{w}$ the mean vector and covariance matrix of the output process are given by $\mathbf{m}_{\boldsymbol{x}} = \mathbf{A}\mathbf{m}_{\boldsymbol{w}}$ and $\mathbf{C}_{\boldsymbol{x}} = \mathbf{A}\mathbf{C}_{\boldsymbol{w}}\mathbf{A}^T$. Thus

$$\mathbf{m}_{\boldsymbol{x}} = \mathbf{0}$$

and

$$\mathbf{C}_{\boldsymbol{x}} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \rho & 1 & 0 & \cdots & 0 \\ \rho^2 & \rho & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^n & \rho^{n-1} & & \cdots & 1 \end{bmatrix} \begin{bmatrix} \sigma_{\mathrm{o}}^2 & 0 & 0 & \cdots & 0 \\ 0 & \sigma_{\mathrm{o}}^2 & 0 & \cdots & 0 \\ 0 & 0 & \sigma_{\mathrm{o}}^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sigma_{\mathrm{o}}^2 \end{bmatrix} \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^n \\ 0 & 1 & \rho & \cdots & \rho^{n-1} \\ 0 & 0 & 1 & \cdots & \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

After some algebraic manipulation this last expression can be put in the form

$$\mathbf{C}_{\boldsymbol{x}} = \frac{\sigma_{\mathrm{o}}^2}{1-\rho^2} \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^n \\ \rho & 1 & \rho & \cdots & \rho^{n-1} \\ \rho^2 & \rho & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \rho \\ \rho^n & \rho^{n-1} & \cdots & \rho & 1 \end{bmatrix} - \frac{\sigma_{\mathrm{o}}^2\rho^2}{1-\rho^2} \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^n \\ \rho & \rho^2 & \rho^3 & \cdots & \rho^{n+1} \\ \rho^2 & \rho^3 & \rho^4 & \cdots & \rho^{n+2} \\ \vdots & \vdots & \vdots & & \vdots \\ \rho^n & \rho^{n+1} & \rho^{n+2} & \cdots & \rho^{2n} \end{bmatrix}$$

The first matrix has the interesting property that the elements on each principal diagonal are equal. It will be seen later that this type of covariance matrix, known as a *Toeplitz* matrix, arises whenever a random process is stationary.

The random process $x[n]$ is thus seen to have a stationary component that arises from applying the stationary sequence $w[n]$ to a linear *time-invariant* system. The second matrix is a *Hankel* matrix and has the property that all elements on the reverse diagonals are equal. This component of the covariance represents the 'transient response' for the system resulting from applying the noise at $n_{\rm o} = 0$. If $|\rho|$ is less than one, which is the condition for stability of the linear shift-invariant system, this transient component eventually disappears (one can observe that the terms in the lower right block of this matrix get closer and closer to zero). In the limit when the output sequence is observed at some time far removed from when the input was first applied, the resulting covariance has only the stationary (Toeplitz) component.

The transient portion of the covariance matrix can be eliminated completely by changing the initial variance of the white noise (i.e., at $n = n_{\rm o}$) to the value

$$\sigma_{\rm o}'^2 = \frac{\sigma_{\rm o}^2}{1-\rho^2} = \sigma_{\rm o}^2 + \frac{\sigma_{\rm o}^2\rho^2}{1-\rho^2}$$

The white noise covariance matrix then becomes

$$\mathbf{C}'_{\boldsymbol{w}} = \begin{bmatrix} \sigma_{\mathrm{o}}'^2 & 0 & \cdots & 0 \\ 0 & \sigma_{\mathrm{o}}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{\mathrm{o}}^2 \end{bmatrix}$$

and the product $\mathbf{A}\mathbf{C}'_{\boldsymbol{w}}\mathbf{A}^T$ contains an extra term that cancels the transient part of $\mathbf{C}_{\boldsymbol{x}}$. The transient portion can be thought of as arising from a mismatch between the steady state variance of the output process and that of the initial input when we force $\mathrm{Var}\left[x[n_{\mathrm{o}}]\right]$ to be equal to $\mathrm{Var}\left[w[n_{\mathrm{o}}]\right]$. Changing the the variance of $w[n_{\mathrm{o}}]$ to match that of the steady state response eliminates the transient. $\square$